IS SYMPATHY NAIVE?

DAI ZHEN ON THE USE OF *SHU* TO TRACK WELL-BEING

(Pre-print. The final version is published in *Taking Confucian Ethics Seriously: Contemporary Theories and Applications*, Julia TAO, Philip J. Ivanhoe, and YU Kam Por, eds., pp. 145-162)

JUSTIN TIWALD*

The mid-Qing philosopher Dai Zhen 戴震 (1724–1777) is famous for his criticisms of orthodox Neo-Confucianism, especially of the Cheng-Zhu 程朱 School that had, by his time, prevailed over intellectual life and state institutions for several centuries.[1] The heart of his critique rests on a controversial series of claims about the Confucian emotional attitude of sympathetic understanding (*shu* 恕) that he found to be essential for moral virtue.[2] As Dai sees his Neo-Confucian adversaries, their account of moral agency puts a stranglehold on sympathetic understanding, for their strictures against the use of desire in moral reasoning prevents someone from sympathetically appreciating other human beings in the requisite ways. Without having healthy desires of one's own, and a sufficient understanding of what those desires should be, we fail to discern the standards of good order (*li* 理) inherent in our condition.

This much is routinely observed. What is less appreciated, however, is that Dai's claims about the importance of sympathy and the desires in moral reasoning are parasitic on a claim about the importance of sympathy and the desires in reasoning about human welfare or well-being. We depend on them to know the standards of good order because, in large part, we depend on them to know what is *good for* people. The

* JUSTIN TIWALD, Assistant Professor, Department of Philosophy, San Francisco State University. Specialties: Chinese philosophy, ethics, political philosophy. E-mail: jtiwald@sfsu.edu

importance of this to Dai's critique is attested by his tendency to portray it as an insight he recovered from the ancients, after it was lost on the Song-Ming Neo-Confucians and their adherents.[3] It is also attested by some of his most plaintive statements, repeated like a refrain, describing the state of blindness brought about by his Neo-Confucian predecessors' impoverished picture of moral deliberation. Without the desires and *shu*, he laments again and again, we preclude ourselves from appreciating the ways in which our general moral guidelines or ideals can become detached from real-world avenues of benefit and harm and thus find ourselves *unknowingly* "bringing irreparable harm (*huo* 禍) to all under Heaven."[4] It is this particular function of sympathetic understanding, where it serves as a way of tracking benefits and harms that I shall focus on here.

Most of us believe that sympathy figures prominently in helping us to better appreciate benefits and harms, and I find the same unspoken assumption in much discussion of *shu.* However, I also find that the mechanism we usually imagine sympathy to rely on for tracking well-being tends to greatly oversimplify and thereby underestimate its importance.

Consider, for example, a case in which a powerful mayor wants to raze an entire neighborhood to make room for airport expansion but, upon realizing some sympathetic appreciation for the residents of that neighborhood, is led to look elsewhere for real estate. How, in particular, has sympathy brought about this change of heart? Normally we assume that it is the mayor's act of imagining himself in the place of the residents (an act I shall call "perspective-taking") that enables him to sympathetically appreciate the extent of the harm he might do.[5] Through sympathy, he imagines what it would be like to lose the home in which multiple generations of one's family was raised, and he focuses on the profound psychic injury it does to deprive someone of a lifelong

community. The problem with this picture, however, is that for purposes of appraising benefits and harms, the affects and interests we imaginatively attribute to others are not always the right ones. For example, we might attribute to others desires they do not have, or they might well have desires that are self-destructive. The response, then, is usually to correct for the desires that we imaginatively simulate, and to do so by specifying that they should be rational or *informed* desires—those that we would have with working faculties and under full information.

This is the move I find to be, at one and the same time, missing from Dai's account and profoundly unfair to sympathy's robust role in moral deliberation. It is unfair to sympathy because it tends to slough off the hard work of determining which desires are the right ones onto something else, when in fact it is sympathy itself that does a great deal of this work. This move is missing from Dai's account because it tends to portray the final arbiter of benefits and harms as a kind of idealized first person, deliberating (rationally and with full information) about what she wants. For Dai Zhen, as I illustrate, wanting something alone is not enough to make it good for us. It must also be something others would want *for* us, insofar as they sympathetically understand us. I find in Dai Zhen a more robust and plausible account of *shu* than mere perspective-taking allows, and this chapter is devoted to the explication of such an account.

A word about terminology. In making this argument, I refer to the ability to track benefits and harms as an ability to track "welfare" or "well-being." I use these latter terms in a thin sense, which describes whatever it is that improves or declines when one does well or does poorly. The advantage of using it in this way is that it helps to unify Dai Zhen's many statements to the effect that the more orthodox picture of moral agency prevents its adherents from taking account of various forms of harm in

their moral deliberations. Whether a particular iteration of this refrain invokes the "harming of others" (*huo ren* 禍人),[6] or more dramatically the "injury" (*shang* 傷) done to "the people" (*min* 民),[7] the point is the same: without making use of our own desires and *shu*, we have no way to guarantee that our moral judgments will take account of the well-being of the affected parties.

To be sure, there is no single word in Dai Zhen's philosophical lexicon that corresponds to "welfare" or "well-being." However, it would be a mistake to conclude from this that there is nothing systematic to say about human well-being and its function in Dai's moral thought. Indeed, if we look almost anywhere in the history of philosophy, well-being tends to be the one normative concept most conspicuous in its lexical absence, for moral thinkers of all ages have rarely written about well-being as such and have instead assumed (much like an ordinary speaker today) that their audiences will understand implicitly that references to such things as "happiness" or "prosperity" all have some bearing on a common thing—well-being—and furthermore that this common thing is important for ethics.[8] Indeed, we should be worried if it were not.

How does a sympathetic understanding of others bring their well-being to bear on our moral evaluations? Here I suspect most of us have a ready answer: by allowing us to reconstruct the *point of view* of others, allowing us to experience it (or some simulation of it) for ourselves. Such is the nature of this sort of exercise, however, that it cannot possibly be a dependable indicator of well-being on its own. Consider the particular kind of perspective-taking advocated by Dai Zhen, where we track the good of others, in part, by asking what we would "desire" (*yu* 欲) if we were them.[9] This might work well enough in certain paradigmatic cases, when weighing the important and nearly universal desires for companionship, nourishment, a stable source of income,

etc. But surely there are times when the moral agent's desires are simply the wrong ones to use in the synthesis of the other's point of view. I mean this in at least two senses. First, they might be "wrong" in the sense that the other may not have them, thus making the object of desire less beneficial for the other than it is for the moral agent. I should always bear in mind, for example, that not everyone is fond of keeping cats as pets, as I am.[10] Second, they might simply be "wrong" in the sense that they have little relationship to well-being, whether or not they rightly inform our reconstruction of the other's point of view. I might legitimately share with Wang a powerful yearning to see our home team win. But it would be strange to say that achieving the object of our desire (the triumph of our home team) contributes to Wang's well-being in the same way that having a constant source of nourishment does, even if she desires it with the same intensity. Just because Wang has the desire, in other words, it does not follow that satisfying it will contribute to her welfare.

If we want to preserve the perspective-taking contributions of sympathy, then we will have to prop them up with further insights and specifications. We will have to find some way of making those desires that have an intimate relationship to the other's well-being, and only those desires, normative for our sympathetic reconstruction of her point of view, so that we do not mistakenly project a strong desire for house cats on an ailurophobe, or an unhealthy obsession with winning on anyone. Something must aid our exercise in perspective-taking so as to guarantee that we extend the right sorts of desires to others, and the key test of those desires, surely, will be whether they do an adequate job of tracking the ways in which the other can be benefited or harmed—that is, ways in which circumstances will affect the other's welfare.

This is surely in the spirit of Dai Zhen's project. Dai does not think that just *any* desire we happen to have should contribute to our sympathetic reconstructions of

another's point of view. When Dai refers to the use of "desires" in moral deliberation, he means specifically those desires that belong to the "ordinary feelings of human beings" (*ren zhi changqing* 人之常情).[11] These include the desires that belong to us by nature and have a universal or near-universal status in ordinary human beings.[12] They are also explanatorily basic desires—that is, desires in terms of which we can properly explain why one wants something more specific or idiosyncratic. Thus, the desires we should have in ourselves are not for things like a longing to share a plate of fettuccini with one's beloved, but a hunger for food (*shi* 食) or a yearning for romantic love (*nannü* 男女).[13]

Most of us believe that our desires provide us with crucial insight into well-being, and that they do so in part because our good is importantly related to desire satisfaction. Part of what makes lifelong companionship a good for me, it seems, is that (at least under certain idealized circumstances) I want it. Dai Zhen shares these views deeply and emphatically, but it is not clear that he shares it for the reasons we might expect. Generally speaking, we could give two sorts of accounts of how desire satisfaction is related to well-being. One would be to see desire satisfaction as beneficial by virtue of some independent value, such as the pleasure or happiness that results from attaining one's desired ends. We cannot take much joy in life, we might think, unless we sometimes meet our felt goals and ambitions; thus, desire satisfaction is useful instrumentally, to facilitate joy. If there is an obvious alternative to the view that desire satisfaction is good by virtue of some independent value, it is that desire satisfaction *simply is* the good. That is, to satisfy someone's desires is just what it means to benefit that person.

Of these two general strategies, certainly one of the prevailing temptations is to read Dai as advocating the second. Dai Zhen identifies the sage-kings' concern about

their subjects' states of well-being (identified in classical sources by such terms as "hardship" [*kun* 困] and "poverty" [*qiong* 窮]) with their attendance to their subjects' desires. Furthermore, desire fulfillment (*sui yu* 遂欲) forms an important leitmotif in Dai's moral philosophy, being invoked again and again as having a much-neglected explanatory power in proper moral judgment.[14]

There is considerable evidence to suggest, however, that if anything Dai prefers a variant of the first of the two ways of relating desire to well-being, where the fulfillment of desires gets its worth from some other more fundamental and independent value. In Dai's case, however, the likely candidate for the latter value would not be happiness or pleasure as such, but would almost certainly be "life fulfillment" (*sui sheng* 遂生). "In human existence," Dai declares, "there is no greater affliction than to lack the means to fulfill one's life (*sui qi sheng*)."[15] Life, for Dai, brings with it its own set of demands.[16] That is, just by virtue of being living creatures, we all have needs for certain goods, such as sustenance and development. The needs that attach to life as such then take different forms according to the inborn "nature" (*xing* 性) of the living thing in question.[17] And it is to the particular desires that arise from the structural requirements of living (and, we might add, "growing") that the truly virtuous or "humane" (*ren* 仁) person is supposed to attend: "If in desiring to fulfill one's own life, one thereby fulfills the lives of others, this is humanity."[18]

Conceiving welfare as "life fulfillment," in turn, allows Dai to link it more directly with the supreme good (*shan* 善) of sustaining the universe's generative or life-producing processes (*sheng sheng* 生生).[19] "Life" (*sheng*) becomes a good when put into harmonious action with the productive forces of Heaven and Earth. This is not to say that it is intrinsically good: to be valuable it must cohere with natural forces, as a

working part of the whole.[20] But life also has demands of its own, being inherently dynamic and productive, and in these demands we find our grounds for pursuing our own well-being. In other words, life fulfillment has value prior to and partially independent of our desiring it, and it is the high price we put on the life in ourselves that inspires us to act in our own interest: "Because all creatures of blood and breath know the love of life and fear of death, *therefore* they pursue benefit and avoid harm."[21]

The question we must ask then is how someone like Dai Zhen can maintain both that *shu* (or "sympathetic understanding") helps us to track the well-being of others by projecting certain desires onto them and at the same time that it does not commit us to endorsing the judgments informed by whatever desires we happen to have. How can we be more selective about the sorts of desires we imagine the other having, so as to give the life-fulfilling ones more normative weight? Here I think we can mention two general answers. First, we might imagine that independent constraints are imposed on the desires selected by *shu.* Second, we might imagine that *shu itself* is selective about the desires that it causes us to synthesize. As I argue, Dai thinks the first sort of answer is helpful to a degree, but he is unusual in stressing the second answer as well.

In the spirit of the first answer, one might point out that we are not always obliged to endorse the picture of the other's psychological landscape that our sympathetic imaginations are inclined to create.[22] We might imagine a fellow fanatical fan being willing to sacrifice his career to see the home team win, but on reflection allow that this preference is simply too strong or irrational. Another appealing strategy is to regulate the sorts of desires that go into our deliberations in the first place. I would be less inclined to attribute wild or self-destructive desires to Wang if I myself am a person of moderate temperament, or at least if I can recognize which of my desires would be inappropriate. Dai's philosophical antagonist Zhu Xi 朱熹 (1130–1200)

tackles a perceived problem in mistaken desires in just this way, suggesting that we first see to it that we clearly understand the standards of moral order (*ming li* 明理) and rectify our hearts and minds (*zheng xin* 正心) and only then proceed to use ourselves as the standards by which we measure others.[23]

Dai Zhen certainly believes we should avail ourselves of these sorts of correctives. A crucial component of being a good judge of others' inclinations, for example, is having at least reasonably well-ordered inclinations oneself. And Dai allows that we can develop a knack for second-guessing our sympathetic judgments over time.[24] But it is much less certain that Dai would share the intuitive grounds for adopting these strategies as I have described them above. To motivate these strategies, after all, I simply assumed that some check against mistaken desires had to be introduced either before the exercise of *shu* (in the corrective inspired by Zhu Xi) or after it (in pointing out that we could refuse to endorse our *shu*-based judgments). But this omits another possibility that, when compared with our everyday processes of moral deliberation, should strike us as the far more natural one: namely, that *shu* itself plays some critical role in determining which sorts of desires are true indicators of our well-being and which are not.

We can appreciate the importance of this latter alternative by comparing it with the model inspired by Zhu Xi.[25] By Zhu Xi's account we begin either with refined desires in ourselves, or at least by specifying which desires would qualify as refined, and once we have done that we go on to practice *shu*. But this fits awkwardly with our everyday experience of sympathy-based moral reasoning (which *shu* may well not be, on Zhu's reading). If some exigency tempts me to miss a lunch with George, and I want to imagine how George might feel about this before committing to it, it would be strange to say that I should have either perfectly "rectified" desires, or a clear account

of which desires are the right ones, in advance of putting myself in his shoes. A major motivation for Dai's turn to *shu* is his supposition that it can help less-than-ideal moral thinkers muddle through their daily lives,[26] and for these purposes Zhu's prerequisites seem too rigorous to put into practice. It makes much more sense to say that my attempt to sympathetically reconstruct his point of view itself plays some selective role, helping me to determine which sorts of desires I should take into account as I proceed. Much as grief for the passing of a loved one, for example, lays bare the ways in which the deceased contributed to our personal fulfillment—ways of which we often had little or no awareness prior to grieving—so too does sympathy help to focus and make more vivid the facts that are salient for determinations of well-being.

If we can accept this as a psychological possibility, then this presents us with two distinct models of sympathetic understanding. Under the first model, *shu*'s real work in tracking well-being is accomplished through perspective-taking. Its primary contribution to my attempt to understand George, and know what is good for him, lies in my imagining myself as him, feeling as he would feel and wanting as he would want. *Shu* is "naive" in the sense that it adopts a set of desires uncritically. This is not to say that *we* adopt them uncritically—as we saw, this naive model is perfectly compatible with the view that we should exercise discretion either before or after our sympathetic reconstruction of the other's point of view. But it does imply that *shu* exercises no discretion itself, and this is precisely what the second model rejects. For the latter, which I shall ultimately attribute to Dai Zhen, *shu* plays a crucial part in helping us to determine which sorts of desires are true indicators of well-being, because *shu* is not just an exercise in perspective-taking, but also a way of valuing a person.

In pursuit of unearthing this more robust account of *shu* in Dai Zhen's work, let me begin by offering a diagnosis of the motivations for the naive model of *shu*, where

its contribution to estimations of welfare lie exclusively in perspective-taking. One of the chief reasons that this model is so alluring is that we moderns (especially those of us who are steeped in modern philosophy and social science) are predisposed to lend a certain explanatory authority to first-person deliberations about one's own desires.[27] If I want to say why a particular apple is good for Mary, I will be inclined to tell a story that ends in claims about what Mary wants. I will most likely explain that she has a hunger for this particular kind of apple, and that the sustenance it offers will help her to get other things that she desires. If I cite benefits that do not immediately invoke her desires—for example, if I mention that having an apple will help to sustain her concentration through the next few hours of work—I am likely to explain the goodness of these other things in terms of more fundamental desires (such as the desire to do good work, to win a promotion, etc.). To be sure, it will not always be the case that Mary wants the apple, for it is not always the case that people want what is best for themselves. But we tend to correct for this problem by saying that the apple is something she *would* want if only she were apprised of all of the relevant information and capable of making sound inferences from that information. If only she knew that eating it would improve her concentration, and if only she knew how it tasted, she would then desire it.

So understood, this deference to informed calculations about one's own desires helps to explain why we tend to believe perspective-taking does all of the relevant work. After all, to determine what constitutes Mary's well-being, it is sufficient to know what she would (under epistemically ideal circumstances) want. If I err in determining what she would want, this is because I make bad inferences, or I do not have access to all of the relevant facts (for example, that she has an allergy to apples). But these are mistakes in reasoning or information gathering, not mistakes in sympathizing itself. Thus, the

explanatory power attributed to Mary's informed desires has an important epistemic implication as well. My inquiry into the constituents of her well-being may safely stop at the point that I am able to determine what Mary would ideally want. Sympathizing with Mary may require more than this kind of perspective-taking, but insofar as we rely on sympathy to know what's good for her, this perspective-taking exhausts its epistemic contribution.

For Dai Zhen, both the naive account and its motivating assumptions would be highly dubious. The first and perhaps most notable problem is that it commits us to a peculiar doctrine about altruistic or disinterested desires: namely, that they are always mistaken desires, and that it would be impossible to rationally want something self-sacrificial (or at least not beneficial) in light of full information. Since the informed desire theory holds that my good is whatever I would rationally want given full knowledge of a certain kind, it is committed to the view that I could never under those circumstances want something bad for myself. This is a major criticism of "informed desire" or "full information" theories of welfare in contemporary philosophy and social science, a criticism that Amartya Sen has captured succinctly in accusing informed desire theorists of "definitional egoism."[28] And it is not difficult to see how it would run against the grain of Dai Zhen's deeply Confucian understanding of human moral psychology.

Consider, for example, the desire of parents for their children's future health and prosperity. It is reasonable to assume, surely, that many mothers and fathers would in light of full information quite rationally want things for their children that come at great cost to themselves. Many are willing to give up a great deal of their freedom and leisure to see to it that their children have the kind of lives that, long after their parental benefactors are gone, the children will continue to find meaningful and satisfying.[29]

And this sacrificial urge to provide for the long-term well-being of one's children tends to be both deeply felt and something of which its possessors are intimately aware. If this sort of thing does not qualify as a desire, it is difficult to imagine what would.

Defenders of the informed desire approach might try to explain this phenomenon away by suggesting that in fulfilling the interests of one's children, the benevolent parent also fulfills her own interests—that the utility function of the parent reflects the utility function of the child, as economists sometimes put it. But this is too facile an explanation. The very language of this sort of sacrifice requires that we allow for trade-offs between parent and child welfare. One might even feel a powerful drive to give up her life in order to spare her child from pain, and while it might give her some comfort to know the benefits of her sacrifice as she goes to her grave, it would be both wrong and unfair to describe such comfort as a good for her that is proportional to the strength of her desire for it.

Given the place of prominence of filial piety in the Confucian moral order, Dai tends to stress the particular subset of desires where the sacrificial relationship is reversed, so as to highlight the filial disposition to want the well-being of one's parents over and above one's own. This disposition exists not just in human beings, but (Dai claims) in any living creature that has an awareness of its parents and thus can be seen even in the way birds feed their mothers in old age.[30] But despite this emphasis, Dai is clear that many kinds of self-sacrificial inclinations exist, not only for one's parents, but also for one's progeny, one's mate, and even in some minimal sense for members of one's own species.[31]

These sorts of phenomena are ones that any reasonable picture of human psychology must accommodate, and they point to an entire sphere of informed desires that do not fall neatly within the purview of self-beneficial ones. More to the point,

however, is that they point to a more nuanced account of self-interest (*si* 私). Consider what sense (or nonsense) we would have to make of the claim that Martin, who is fully informed about the choices before him, "wants his own good." For an informed desire thinker this sort of statement would be redundant: the fact that Martin wants *anything* under these epistemically ideal circumstances is sufficient reason to take it to be good for him. But I suspect that most of us will find this to run against our considered convictions, for we tend to understand self-interest not just as wanting whatever we would want under full information, but also of having a certain self-directed pro-attitude. To be self-interested is to consider one's own well-being valuable and thus to desire it as an end.[32] This, surely, is a noticeable feature of Dai Zhen's understanding of self-interestedness:

> Whether one cares self-interestedly only for oneself (私於身), or whether one extends [this care] to those near and dear to oneself, these are both kinds of humane love. To care for oneself is to love oneself (仁其身), and to extend it to those near and dear to oneself is to love one's intimates.[33]

There are a few things to note about this passage. First, if Dai thought that informed desires were definitionally egoistic, it would be strange to add that self-interest (*si*) requires the attitude of "loving" (or "loving humanely" [*ren* 仁]) oneself. This sort of attitude, understood by most Confucian thinkers as requiring a substantial mechanism of psychological habits and attitudes, would appear to be the epiphenomenal icing on the motivational cake. Having correct desires would already be enough.

Second, Dai's point here is that self-interest and generous feelings toward one's intimates are not really different in kind. Both share a fundamental feature in common. What is it? If we look at the way Dai sets up this proclamation about self-interest, we can begin to see the contours of an answer:

> All creatures of blood and *qi* know the love of life and fear of death, and therefore pursue benefit and avoid harm. Although they differ in understanding, they are nevertheless the same in not going beyond this love of life and fear of death…. The love of that which has given one life [one's parents], the love of that to which one has given life [one's children]…all of this proceeds from the love of life and fear of death.[34]

Earlier I suggested that Dai Zhen's account of well-being is best understood as what he calls "life fulfillment" (*sui sheng*), a turn of phrase deeply evocative of Dai's highest good, "producing and sustaining life" (*sheng sheng*). Here again Dai makes the connection between life and well-being explicit: insofar as we love life and fear death, he says, we seek benefits and try to avoid harms. The question, however, is whose benefits we seek and whose harms we avoid, and Dai's answer is that it depends on whose life we love. Whether we want our parents to fare well or whether we want ourselves to prosper, both sorts of want share a common attitude: namely, a love of the life that belongs to the person in question. This suggests, *pace* the informed desire view, that knowledgeable desiring is not, ultimately, a sufficient condition for self-interest. To be self-interested we must also have the right sort of feeling—"love" (*huai*

懷)—toward our own life, with all of the structural demands of which that life is constituted.

To be sure, it is much easier to elicit a love of one's own life than it is to love another's. Dai seems to assume that for most of us the latter will never come so effortlessly as the former, which is one reason why the imaginative exercise of putting oneself in another's place is so important.[35] But Dai reminds us that just because our self-love comes relatively effortlessly we should not assume that this feeling, so crucial to wanting the good of someone else, is somehow missing in cases where the life in question is one's own. To care deeply about one's own well-being requires the same robust feeling of attachment as caring deeply about the well-being of anyone else.[36]

The final point to make about the passage on *si* is that it has implications not just for the motivational requirements of self-interest, but for the epistemic requirements as well. That is, we depend on a self-directed sense of humanity even to understand what is in our own self-interest. To be humanely disposed toward someone is more than a matter of being inclined to act on a series of already specified ends, as though it could be our humanity that drives us to help a child, but something else that specifies the particular ways in which the child needs help. Humanity is also the virtue by which we come to recognize which ends are worthwhile in the first place.[37]

This helps to fill out the more robust, nonnaive account of *shu* that we need. It suggests that this felt attachment to someone's life and life fulfillment, over and above the attachment to someone's desire fulfillment, is doing some of the necessary work in helping us to distinguish the truly self-interested desires from the rest, even in the cases where the desires in question are our own. There is remarkably strong prima facie evidence for this more nuanced phenomenology of self-interest. Surely we can distinguish between "what we want" and "what we want for our own sake," at least

much of the time. I know that having my home team win would not benefit me in proportion to the strength of my desire for it, and it seems likely that my ability to adopt some caring or loving stance toward myself has something to do with this insight.[38] Needless to say it is not *sufficient* for this insight—and Dai Zhen is eager to supplement our deliberations with the refinement in judgment and knowledge of the world that comes with effective self-cultivation—but it is *necessary* all the same. Most of our attempts to take the well-being of others into account requires this "love of life and fear of death" at a minimum. It is an intellectualist's fantasy that we will appreciate the relevant demands of life with sufficient depth, that we will be able to see them vividly, without this love.[39]

In the fundamental matter of taking proper account of welfare goods then I find this "love of life"—so often invoked by Dai—to be absolutely vital. However, there is a second contribution made by "humanity toward oneself" that I want to offer in a slightly more tentative way, and that contribution is a shift in point of view. For Dai Zhen, the feeling of humanity or benevolence toward someone, which he understands as a feeling aimed at "fulfilling one's life,"[40] almost always contains within it some sympathetic understanding of the people whose lives are to be fulfilled. To be sure, not all close followers of Confucius see humanity as constituted by some exercise of *shu*, and many of Dai Zhen's Neo-Confucian predecessors insist that *shu* should eventually drop out of our deliberative repertoire as we strive to become humane people.[41] But for Dai Zhen *shu* figures centrally in being humane, and generally speaking one cannot be the latter without at the same time exercising the former.[42] It is quite likely, therefore, that self-interest requires "humanity" because the latter includes within it a kind of sympathetic understanding. The act of considering myself from a sympathetic stance—as a concerned observer might consider me, comparing her case to my own—plays an

important part in helping me to highlight the desires whose ends are important for my own life fulfillment.

A revealing detail in this respect is Dai's tendency to describe *shu* not just as a way of studying the feelings of others, but also as manner of "returning to oneself" (*fan gong* 反躬),[43] an expression he takes from the "Record of Music" ("Yueji" 樂記).[44] For Dai Zhen, "returning to oneself" is essentially the form of self-examination that we undertake in order to introspectively evaluate our behavior before we commit to it. So quite naturally, Dai typically presents it as something we should do before imposing ourselves on others. "Whenever one does something to another," he offers, "one should return to oneself and calmly reflect: If another were to do this to me, would I be able to bear it?"[45] This may seem to suggest that we use it primarily to gain insight into the would-be psychological landscape of others, but if we scrutinize Dai's analysis of the expression more carefully, we see that it does more work than this. Additionally, it also helps us to distinguish between *our own* desires, telling us which ones are superfluous or meddlesome and which ones really count. And those that really count are the core of basic desires that we share in common with other human beings. This is a notion that Dai takes once again from the "Record of Music," noting how it attributes to all of us a common set of essential dispositions or true feelings (*qing* 情) to which belong the desires that are most essentially our own.[46] Thus, the process of "returning to oneself" accomplishes at least two things: it helps us to understand how we would feel were we in another's shoes, but importantly it also helps to clarify which desires we are rightly expected to take into account. Surely the latter accomplishment sheds as much light on ourselves as it does on others.

"When certain doctrines enter deeply into the hearts of human beings," Dai declares, "their harm is great, and yet no one is able to awaken to them."[47] Dai is clear

that he sees both a healthy use of one's own desires, as well as a proper reliance on *shu*, as the remedy for the prudential blindness that he finds stored away in the moral ideals of his age. What I have tried to do here is show that, in an important sense, the latter that is more fundamental than the former. It is our success at understanding ourselves as objects of sympathy that informs our evaluations of well-being, and not the other way around. *Shu* brings with it the right emotional attachment and sufficient distance from internal point of view to shed light on the sorts of interests and inclinations that truly matter, and so it becomes less the flimsy exercise of perspective-taking and more the familiar emotional attitude of sympathetic understanding as we characteristically see it in our everyday deliberations. As such, it is better suited to serve as the more fundamental source of insight into the human good.

Acknowledgments: My thanks to P. J. Ivanhoe, Julia Tao, and Yu Kam-por for comments on an earlier draft of this essay. I am also indebted to Chris Fraser, Koji Tanaka, and the many other scholars who participated in the Fourteenth Conference of the International Society for Chinese Philosophy, where I first presented a version of the argument I make here.

[1] The chief philosophical works I cite here are Dai's *On the Good* (*Yuanshan* 原善), *Remnants of Words* (*Xuyan* 緒言), and especially *An Evidential Study of the Meaning of Terms of the Mengzi* (*Mengzi ziyi shuzheng* 孟子字義疏證), hereafter *Evidential Study*. A reliable and widely available edition of the two most important of these works, *On the Good* and *Evidential Study*, is attached as a lengthy "appendix" to Hu Shi's 胡適 *The Philosophy of Dai Dongyuan* (*Dai Dongyuan de zhexue* 戴東原的哲學) (Reprint, Taipei: Taiwan shangwu, 1996), 201–39, 240–337. *Remnants of Words* appears in the *Complete Collection of Dai Zhen* (*Dai Zhen quanji* 戴震全集), vol. 1, ed. Dai Zhen Research Group

(Beijing: Qinghua daxue chubanshe, 1991), 64–116. Other primary texts of Dai's will be cited as needed. My somewhat controversial translation of *Xuyan* as *Remnants of Words* (referring to the remnants of *the sages'* words) follows Yu Yingshi, *Lun Dai Zhen yu Zhang Xuecheng* (Hong Kong: Longmen shudian, 1976), 100–101.

[2] *Shu* has been variously understood as the Confucian Golden Rule, the practice of moral reciprocity, and (more often than not) as a distinctively Confucian form of sympathy. Dai himself describes it as "taking oneself and extending it to others" (*Evidential Study*, chap. 15 [*li*/15]), and I have argued elsewhere that this is best understood as encompassing a form of sympathetic understanding, whereby we imagine ourselves by analogy to be in another's place, and so for a time value them like we value ourselves (*Acquiring "Feelings That Do Not Err": Moral Deliberation and the Sympathetic Point of View in the Ethics of Dai Zhen* [Ph.D. diss., University of Chicago, 2006]).

[3] To quote from Dai's closing indictment of the Neo-Confucians in his *Evidential Study*: "When Yao and Shun worried about the 'hardship and poverty within the four seas,' or when King Wen 'looked upon the people as if they were injured,' was there even a single affair in which they were not planning on behalf of the desires of these people?" The wisdom of Wen and the sage-kings, here, lies in the recognition of two facts. First, one cannot know the substantive moral standards of good order (*li*) without also knowing what is good or bad for the would-be beneficiaries of those standards in the first place. To attend to the good, as these ideal figures demonstrate, is to attend to what is identified in classical sources by the terms "hardship" (*kun* 困), "poverty" (*qiong* 窮), and "injury" (*shang* 傷). Later in the same closing statement, Dai says that those who insist on ignoring the desires in their deliberations will bring misfortune or calamity (*huo* 禍) on others (*Evidential Study*, chap. 43 [afterword]). Second, the way we attend to the well-being of

others, and so attempt to overcome hardship, poverty, and other afflictions, is just by "planning on behalf of their desires."

[4] *Evidential Study*, chap. 40 (*quan*/1). See also chaps. 5 (*li*/5), 10 (*li*/10), 41–42 (*quan*/2–3), 43 (afterword), and *passim*.

[5] See, for example, Nancy Sherman, "Empathy and Imagination," in *Midwest Studies in Philosophy* 22 (1998): 82–119, especially her remarks on sympathy (p. 110).

[6] See the end of the final chapter of *Evidential Study*, chap. 43 (afterword).

[7] *Evidential Study*, chap. 43 (afterword).

[8] To understand the basic argument in J. S. Mill's *Utilitarianism*, for example, it is crucial that we be able to identify such turns of phrase as "in the interest of the agent himself," "for the benefit of the individual," or "having full worth" as bearing on the same kind of good (viz. human well-being), and that there are things we can say about this good as a kind—for example, that it should be cashed out in terms of happiness. See Alan Ryan, ed., *Utilitarianism and Other Essays* (New York: Penguin Books, 1987), 272–338, especially 290–98.

[9] Dai's favored account of *shu* comes from its locus classicus in *Analects* 15.24: "What you would not desire yourself, do not inflict upon others" (*Evidential Study*, chap. 5 [*li*/5]).

[10] This sort of problem has been aptly described as a "paternalism" of desire attribution. See, for example, Herbert Fingarette's "Following the 'One Thread' of the *Analects*," *Journal of the American Academy of Religion: Thematic Issue* 47s (September 1979): 373–405, especially 384–85, 388–95; P. J. Ivanhoe's "The 'Golden Rule' in the *Analects*," in *Confucius Now: Contemporary Encounters with the Analects*, ed. David Jones (LaSalle, IL: Open Court, 2007), 81-108.

[11] See Dai's *Evidential Study*, chap. 5 (*li*/5), and his remarks on chapter 13 of the *Doctrine of the Mean*, in the "Supplemental Commentary on the *Doctrine of the Mean*," in *The*

Complete Writings of Dai Zhen, vol. 2, ed. Zhang Dainian (Hefei: Huangshan she, 1994), 47–83, especially 61.

[12] *Evidential Study*, chap. 2 (*li*/2).

[13] Ibid., chap. 21 (*xing*/2).

[14] For example, notice an important passage in chap. 10 (*li*/10) of the *Evidential Study*, where Dai says the sage "governs the empire by embodying the feelings of the people and fulfilling the desires of the people, and so the kingly way is complete." See also chaps. 30 (*cai*/2), 40 (*quan*/1), 43 (afterword), and *passim*.

[15] *Evidential Study*, chap. 10 (*li*/10). See also chap. 40 (*quan*/1).

[16] See Dai's analysis of the Mengzi-Gaozi debates, in which Gaozi maintains that "nature" (*xing* 性) means "life" (*sheng* 生) (*Mengzi* 6A1–6; *Evidential Study*, chap. 21 [*xing*/2]). Following the parameters of this debate, Dai's discussion of "life" is notably confined to things with "consciousness" or "awareness" (*jue* 覺) and therefore omits plant life. On this point see also the "question" section of chap. 27 (*xing*/8).

[17] *Evidential Study*, chap. 31 (*cai*/3).

[18] Ibid., chap. 10 (*li*/10).

[19] More literally, the process of "growing life," a reference to the "Appended Remarks" section of the *Book of Changes* (*Yijing* 易經, part 1, chap. 5).

[20] *On the Good*, 1.1 (other discussions of *sheng sheng* and its moral significance appear in 1.3, 1.4, 2.1, 3.1, and 3.16). Dai, we should note, is not a welfarist. That is, he does not think that the only thing that counts as good, ultimately, is welfare or well-being. He also builds in the requirement that human life maintain a kind of continuity with natural generative processes, and it is actually from this latter sort of good that human welfare derives its moral value (*Evidential Study*, chap. 36 [*ren, yi, li, zhi,* 1]). In this respect, as in many others, Dai's moral thought resonates deeply with Xunzi's. See P. J. Ivanhoe's

"A Happy Symmetry: Xunzi's Ethical Thought," in T. C. Kline III, ed., *Ritual and Religion in the* Xunzi, (Albany: State University of New York Press, 2008), 63-87.

[21]凡血氣之屬皆知懷生畏死，因而趨利避害. Here the Chinese for "therefore" (*yin'er* 因而) subordinates the second clause to the first, so that the second (seeking after profit and avoiding harm) happens under the condition of the first being true (knowing the love of life and fear of death) (*Evidential Study*, chap. 21 [*xing*/2]). In other words, "therefore" should *not* be read in the merely evidentiary sense of "there is smoke therefore there is fire." Rather, it is *because* we love life that we pursue profit, and thus Dai clearly implies that loving one's life is a prerequisite for seeking after one's own good. Dai Zhen makes strikingly similar statements in *On the Good*, 2.2, and *Remnants of Words*, 1.16.

[22] A point made by Nancy Sherman, "Empathy and Imagination," 110–11.

[23] If we do not take these precautions, he says, "then that which we desire will not necessarily be that which we should desire, and that which we detest will not necessarily be that which we should detest. If, without examining these things, one hastily uses his own desires as the standard for that which he imposes on others, then even if his intentions are impartial (*gong* 公) his deeds will be self-serving (*si* 私)" (*Concordance to Master Zhu's "Questions and Answers on the Four Books,"* ed. Toshimizu Goto [Hiroshima: Hiroshima Daigaku Bungakubu, 1955], 487).

[24] *Evidential Study*, chap. 41 (*quan*/2).

[25] I say this model is "inspired by" Zhu Xi, rather than directly attributable to him, because it is uncertain that Zhu is particularly concerned that the right desires be directed toward one's own good. It might be, after all, that he sees it primarily as a matter of making them line up with the moral good. As Zhu is generally a eudaemonistic thinker, I suspect he believes they should line up with both, but his analysis in the *Questions and Answers* (quoted above) is ambiguous.

[26] In chapter 10 of the *Evidential Study*, Dai notes that a great deal of speculation about moral wisdom focuses on the sage figure, and yet that the rest of us still need to call on our own less-than-sagely faculties in order to resolve our everyday affairs: "All people have business that concerns the household, the affairs of state, and the world. Can it really be that we must await the wisdom of the sage before we can act?" I have argued elsewhere that Dai offers *shu* as the preferred form of moral deliberation for imperfect adjudicators like ourselves (*Acquiring "Feelings that Do Not Err,"* chap. 1).

[27] For example, we find this tendency in such luminaries as John Rawls (*Theory of Justice* [Cambridge, MA: Harvard University Press, 1971], chap. 7), R. M. Hare (*Moral Thinking: Its Levels, Method, and Point* [Oxford: Clarendon Press, 1981], chap. 5), and Joseph Raz (*The Morality of Freedom* [Oxford: Clarendon Press, 1986], chap. 12).

[28] Amartya Sen, *Choice, Welfare, and Measurement* (Oxford: Basil Blackwell, 1982), chap. 4.

[29] Mark Overvold has made effective use of a case much like this one in "Self-Interest and the Concept of Self-Sacrifice," *Canadian Journal of Philosophy* 10 (1980): 105–18, especially 108.

[30] Not all animals show the same sort of filial piety, nor do they tend to make the same sorts of sacrifices for their fathers. But Dai says this is because they are unaware of their parentage, not because they lack the relevant dispositions. See *Evidential Study*, chap. 21 (*xing*/2).

[31] All living creatures show some basic willingness to sacrifice for members of their own species, Dai says, insofar as they demonstrate a great reluctance to eat their own kind (*Evidential Study*, chap. 21 [*xing*/2]).

[32] As the ethicist Stephen Darwall points out, a depressive or self-loathing person might not want her own good, and while we might tell her that she is misinformed about what she

wants, there is nothing *conceptually* incoherent in her saying "I know what I really want, and what I really want is bad for me." She may simply fail to value herself (and thus what is good for her) in the way that we would expect (*Welfare and Rational Care* [Princeton: Princeton University Press, 2002], 5–6).

[33] *Evidential Study*, chap. 21 (*xing*/2).

[34] Ibid.

[35] Except, perhaps, in the problematic case of the sage, who is able to extend such love to others. I discuss this likely exception in *Acquiring "Feelings that Do Not Err,"* chap. 2, section 6.

[36] To be sure, it sounds awkward to use language normally reserved for moral praise to describe the essentially self-serving accomplishment of wanting one's own good. And Dai's handling of the phraseology reflects this awkwardness, specifying that the love of one's self and intimates are mere "kinds of humanity" (that is, they are in the territory of humanity or humane love [*ren zhi shu* 仁之屬]). But even if we can only speak in terms of *quasi*-humanity (or *quasi*-humane love) the basic point remains the same: the core attitude that motivates the love of others (a kind of care or concern for a particular person that Dai cashes out as a "love" of that person's "life") is essentially the same attitude that motivates the love of oneself.

[37] *Evidential Study*, chaps. 21 (*xing*/2), 36 (*ren, yi, li, zhi*/1), 41 (*quan*/2), and *passim*.

[38] I owe this crucial observation (and much else in this essay) to Darwall's *Welfare and Rational Care*. I describe some fundamental differences between Dai and Darwall in *Acquiring "Feelings that Do Not Err,"* chap. 5, section 3.

[39] To hark back to an earlier analogy, this is as unlikely as it is that we can appreciate someone's everyday, often invisible contributions without some powerful emotion like grief to highlight them. Grief helps to focus our attention on the little things that the

deceased has done for the griever, making the absence of these things more palpable to the griever, and thus enabling her to recognize and appreciate the deceased's contributions in a way that would never be available to her without her grief.

[40] See the opening lines of *Evidential Study*, chap. 36 (*ren, yi, li, zhi*/1).

[41] Cheng Yi 程頤 (1033–1107), Zhu Xi, and Wang Yangming 王陽明 (1472–1529) all regard *ren* as a virtue not unlike *shu* in its results, in that it makes one do things that harmonize with the desires of others, but as a means they regard *shu* as being too deliberative and affectively laden for truly humane thoughts and behavior. On this point, see David Nivison's "Golden Rule Arguments in Chinese Moral Philosophy," in *The Ways of Confucianism*, ed. Bryan Van Norden (Chicago: Open Court, 1996), 69–70.

[42] Most commentators agree that one can be *shu* (sympathetically understanding) without exhibiting the fuller virtue of *ren*. But they appear to differ on the question of whether one can be *ren* without also at the same time being *shu*. Dai Zhen is clear that at least the nonsages among us cannot and cites Mengzi to make his case: "To act out of robust sympathetic understanding—in seeking humanity, there is nothing closer than this" (強恕而行, 求仁莫近焉; *Mengzi* 7A4, discussed in Dai's *Evidential Study*, chap. 41 [*quan*/2]). See also the last two steps of the three-step ascent to humane virtue, as Dai describes it: "To take one's self and extend it [to others] is 'sympathetic understanding' (*shu*), and to share in the sorrows and joys of others is 'humanity' (*ren*)" (*Evidential Study*, chap. 15 [*li*/15]).

[43] Ibid., chaps. 2 (*li*/2), 11 (*li*/11).

[44] "When one's loves and aversions are not regulated internally, and one's understanding is led astray by external things, one cannot *return to oneself* and the heavenly patterns of good order are extinguished" ("Record of Music," chap. 19 of *Record of Rituals*, *Sibu congkan* edition, 11.7b5–8); emphasis is mine.

[45] *Evidential Study*, chap. 2 (*li*/2).

[46] Here, Dai uses the term *qing* not in the descriptive sense of affective dispositions that we all tend to have, but in the more selective sense of essential traits of character that all of us will have if nurtured in the right way, under conditions of sufficient discipline and self-restraint. I use "*true* feelings" as a shorthand for this latter, more normative sense of the term.

[47] *Evidential Study,* chap. 43 (afterword)